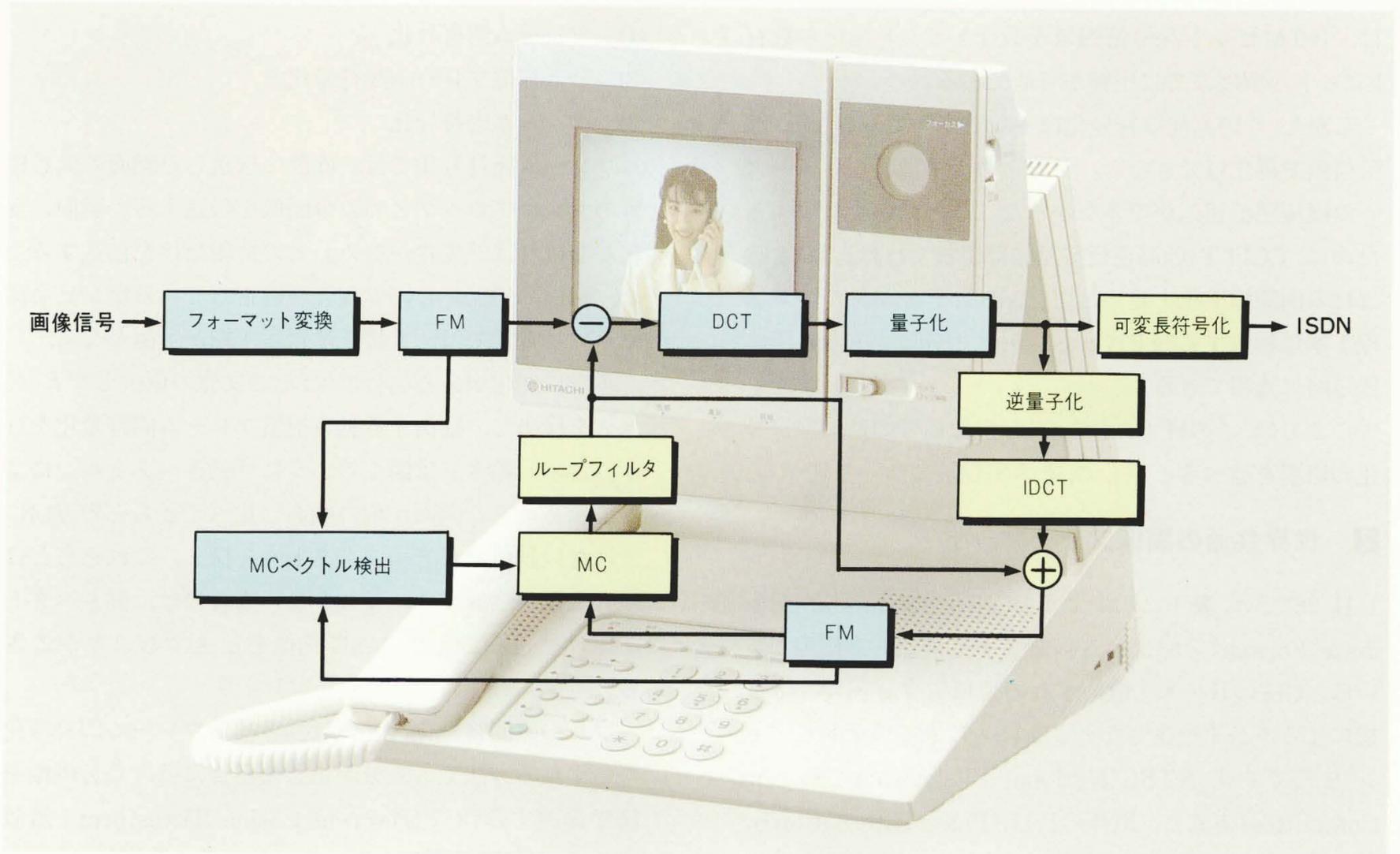


# 情報量の大幅な圧縮を可能とする 画像符号化技術

Video Coding Technology

木村淳一\* Jun'ichi Kimura  
滝沢正明\*\* Masaaki Takizawa  
齊藤 規\*\*\* Tadashi Saitō



注：略語説明 MC (Motion Compensation：動き補償), FM (Frame Memory), DCT (Discrete Cosine Transform), IDCT (Inverse DCT)  
ISDN (Integrated Services Digital Network)

標準化された動画像符号化のブロック図 標準化された符号化方式により、画像を1,000倍以上に圧縮することも可能である。□色の部分だけが標準化で規定される。

テレビ電話・テレビ会議などの画像通信は、長年の夢から出張や会議などによる労力・コストを削減するための重要な道具へと変わってきた。画像信号は音声信号の1,000倍以上の情報量があるため、高能率符号化して大幅な情報量削減を行わなければ簡単に通信することができない。高能率符号化のための国際標準化はCCITT(国際電信電話諮問委員会)によって制定され、MC(動き補償)フレーム間予測やDCT(Discrete Cosine Transform：離散コサイ

ン変換)などの技術が採用されている。しかし、これら符号化技術の複雑さのため装置規模、経済性が問題となっている。

日立製作所では画像信号の高能率符号化方式を分析し、各伝送路に最適なアーキテクチャを開発することにより、小形で経済的なテレビ電話“HV-100”，およびテレビ会議システム“HITVISUAL1500”を開発した。

\* 日立製作所 中央研究所 \*\* 日立製作所 情報通信事業部 \*\*\* 日立製作所 映像メディア研究所

## 1 はじめに

テレビジョン信号に代表される画像信号の信号帯域は非常に広く、電話の音声信号の1,000倍以上になる。幸いなことに画像信号には冗長性が多いため、デジタル的に高能率符号化し、伝送に必要な情報量を大きく削減することができる。テレビ電話やテレビ会議などの用途では、100 Mビット/sの情報量を数十kビット/sから数百kビット/s程度までに圧縮が可能である。

しかし、この高能率符号化は手順が少しでも異なると、受信側で再生はできない。そのため従来は異なるメーカーの機種間で通信ができなかった。この問題を解決するために、CCITT(国際電信電話諮問委員会)はH.261と呼ばれる国際標準化を定めた。H.261は、デジタルの電話線1本に相当する64 kビット/sから2 Mビット/sまでの伝送路に適用できる。

ここでは、このH.261を中心に、画像信号の高能率符号化の概要を述べるとともに、その実現法について述べる。

## 2 世界共通の画像フォーマット

H.261では、表1に示すようなCIF(Common Intermediate Format)と呼ばれる画像フォーマットを使用している。CIFは日・米・欧いずれの地域もフォーマット変換による不公平を生じないように、日米でのテレビジョン方式であるNTSC(National Television System Committee)方式と、欧州のPAL(Phase Alternation by Line)方式の中間が採用された。同時に、CIFの縦横 $\frac{1}{2}$ の大きさであるQCIF(Quarter CIF)も標準化に取り入れられている。わが国では、1フィールド240ラインをデジタルフィルタによって288本に増やしてCIFへの変換を行う。

表1 画像フォーマット H.261では、世界のどのテレビジョン方式とも容易に変換ができるように、NTSCとPALの中間の画像フォーマットを採用している。画面解像度はNTSC, PALの縦横 $\frac{1}{2}$ のものと、 $\frac{1}{4}$ のものと2種類用意している。

		CIF	QCIF	NTSC	PAL	
画面サイズ	(輝度)	水平	352	176	720	720
		垂直	288	144	480	576
	(色)	水平	176	88	360	360
		垂直	144	72	480	576
フレーム周波数(Hz)		29.97	29.97	29.97	25	

## 3 画面の動きを追跡する動き補償フレーム間予測

H.261では、CIFフォーマットの画面を16画素×16ラインのマクロブロックと呼ばれる四角形に分割して、マクロブロック単位に次の三つのいずれかの方法で符号化される。

- (1) フレーム間符号化
- (2) 動き補償フレーム間符号化
- (3) フレーム内符号化

フレーム間符号化では、直前に伝送した画面の同じ位置のマクロブロックとの差分画像を伝送する。画面に変化がなければ「変化がない」との情報だけを伝送するため、画面の50%から70%程度が静止しているテレビ会議やテレビ電話では、大幅な情報量圧縮が可能となる。

画面に変化がある場合には、その変化の原因となる「動き」を検出し、補償する動き補償フレーム間符号化を行う。この「動き」は図1のように「動きベクトル」によって示される。画面が前の画面に比べてどちらの方向にどれだけ動いたかを示すベクトルを探し、これをもとに差分画像を生成する。差分がない場合には、動きベクトルだけを送ることによって画像を伝送することができる。

フレーム間符号化あるいは動き補償フレーム間符号化のいずれの方法でも差分が生じた場合には、差分画像を後で詳述するDCT(Discrete Cosine Transform: 離散コサイン変換)で変換したのちに符号化・伝送する。

また、シーンチェンジなどによって画面の変化が非常に大きい場合は、差分画像ではなく原画像を直接DCTで変換するフレーム内符号化が用いられる。

動きベクトルの検出は、ブロックマッチング法が使われることが多い。ブロックマッチング法は、図2のように原画のマクロブロックと各動きベクトルで示される部分との差分の最も小さいベクトルを採用する方法である。ブロックマッチングは、候補のベクトルの数だけ行う。例えば、ベクトルを水平、垂直それぞれ±15の範囲で探索する場合、候補ベクトルは961個になる。すなわち、画素一つについて961回、1秒間に10億回の演算を行わなくてはならない。動きベクトルの探索を、性能を落とすことなく効率よく行うことが重要になる。

## 4 画像圧縮の主役(DCT)

動き補償フレーム間予測などで生じた画像の差分信号

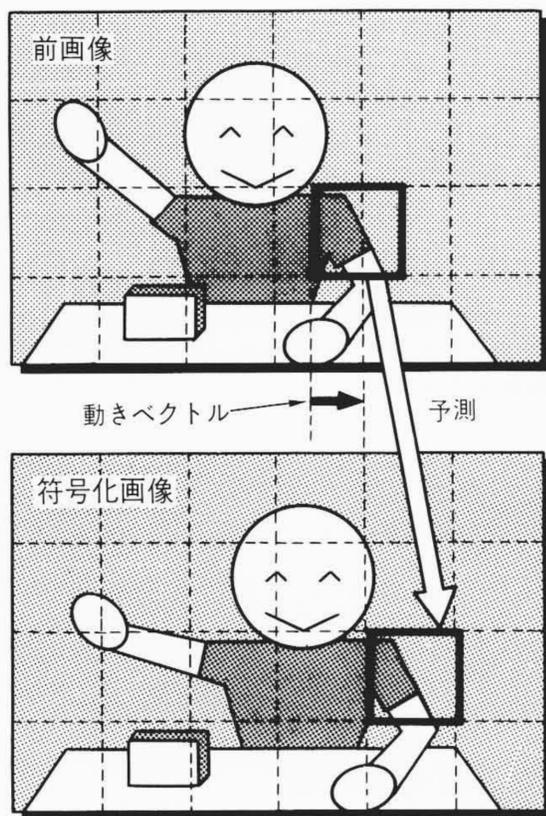


図1 動き補償フレーム間予測 画像が直前に伝送した画像からどのように動いたかを、ベクトルで表現する。例えば、符号化画像の枠内を符号化するとき前画像の枠部分を予測画像として利用する。

は、DCTによって変換し、符号化する。DCTはフーリエ変換などの周波数変換の一種である。H.261でのDCTでは8画素×8ラインの信号を、図3に示す64個のDCT係数と呼ばれる周波数成分に分解する。

画像信号の周波数成分は低周波数(図3の左上方向)に偏って分布しており、しかも高周波数成分に対する人間の視覚感度は弱い。このため、低周波成分だけを符号化することによって、画像信号を効率よく符号化することができる。

DCTの変換例を図4に示す。同図の左の画像信号はDCTによって中央のDCT係数に変換される。DCT係数の各数字は図3の同じ位置の周波数成分の大きさを示す。H.261では、このDCT係数を直流成分からジグザグスキャンと呼ばれる順に読み出し、符号化を行う。

符号化は、ゼロでないDCT係数とその直前に読み出されたゼロのDCT係数の個数を組み合わせて可変長の符号を割り当てる。

画像を符号化して所定の伝送路で伝送するためには、発生する情報量をコントロールする必要がある。その手法の一つにDCT係数の量子化がある。量子化では、DCT係数を量子化ステップサイズと呼ばれる値で割り、その「商」を伝送する。受信側では伝送された「商」に量子化ステップサイズを掛けて元の値を再生する。量子化ステ

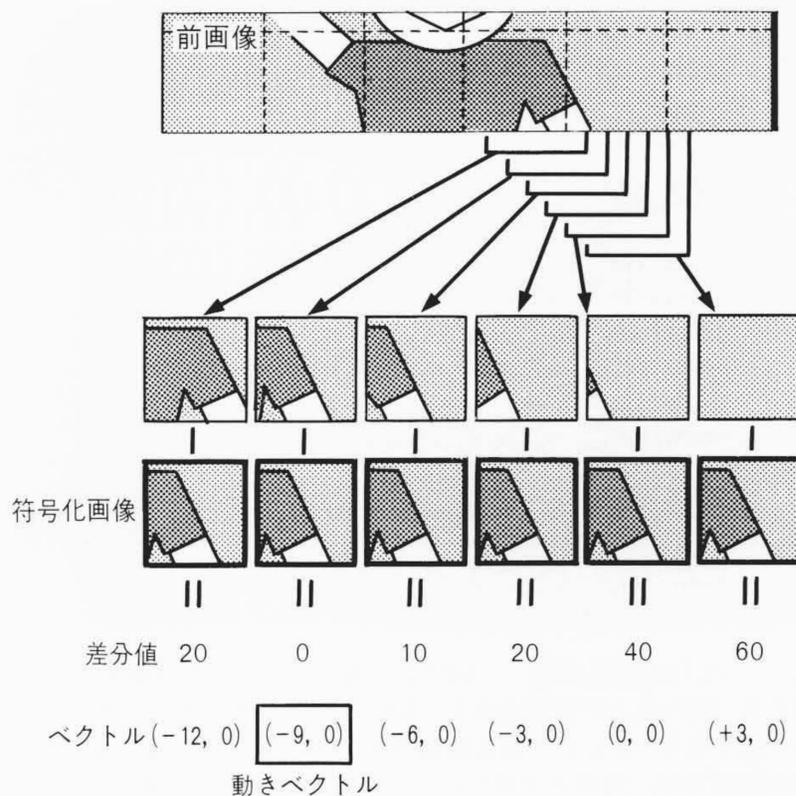


図2 ブロックマッチングによる動きベクトル検出 動きベクトルの検出は、前画像から読み出すブロックの位置を少しずつずらしていき、差分の最も小さい点を求める。差分値は誤差電力などが用いられることが多い。

ップサイズを大きくすれば、発生する情報量が削減すると同時に受信側で得られるDCT係数のひずみが大きくなり画像が劣化する。量子化の例を図5に示す。情報量が少なくなるほど画像が劣化しているようすがわかる。

発生する情報量が増えると、1秒間に伝送できるフレーム数が減少する。画像の劣化と情報量の増加による伝送フレーム数の減少の比率をコントロールすることがテレビ電話やテレビ会議の性能を左右する。

### 5 H.261アルゴリズムの実現法

H.261の処理の全体ブロック図を33ページの図に示す。図の淡いグリーン色を付けた部分が標準化で制限される部分である。これ以外の部分は装置の性能、規模、価格などによって自由に設定できる部分である。これを実現するには何通りも方法があるが、大きく分けるとソフトウェアを主体とする方法と、ハードウェアを主体とする方法になる。64kビット/s程度の低速度のテレビ電話の場合は、処理をやや簡略化してソフトウェア主体の処理をすることによって小形化・経済化を図ることができる。逆に、1.5Mビット/s程度までの回線を使用するテレビ会議は、ハードウェア主体の処理によって高性能化を図ることができる。

テレビ電話“HV-100”のブロック図を図6に示す。

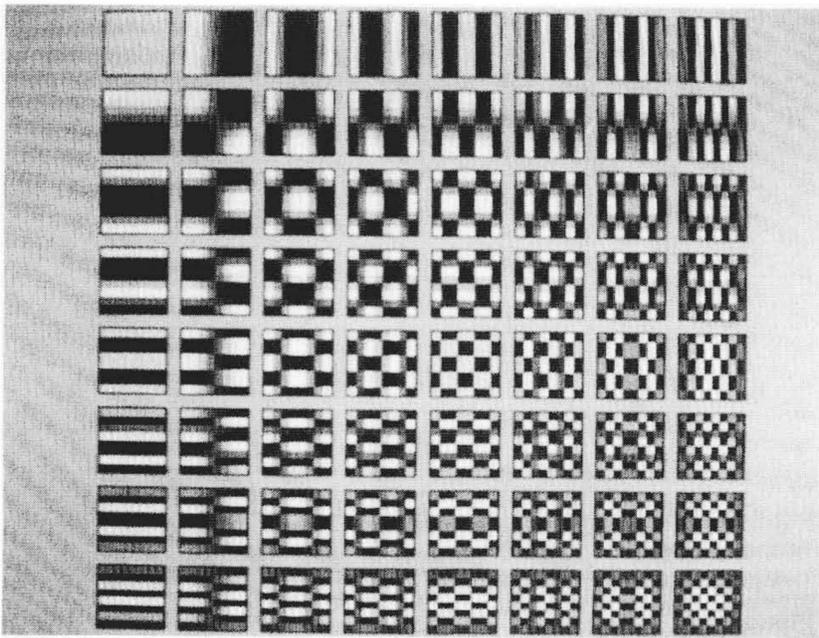


図3 DCTの基底波形 DCTでは画像を64個の周波数成分に分解する。図の左上が直流成分で、右にいくほど水平方向の周波数成分が高くなり、下にいくほど垂直方向の周波数成分が高くなる。通常の画像では、情報の大部分は左上方の低周波領域に集中する。

HV-100ではH.261の処理を、DSP(Digital Signal Processor)によるソフトウェア処理を中心に、一部をゲートアレーによるハード処理で実現している。回線を128 kビット/sに限定し、ソフトウェアとハードウェアの分担を最適化することにより、小形化と経済化を両立させている。

H.261をハードウェアで実現した例として、テレビ会

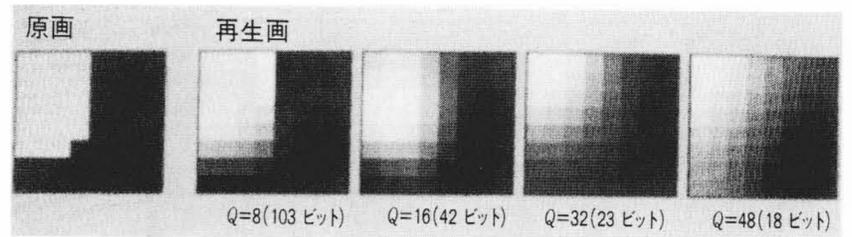


図5 量子化による情報量制御と画質の劣化 量子化ステップサイズ(Q)を変化させることによって、画像の情報量の削減や高画質化を図ることができる。量子化ステップサイズの決め方が符号化装置の性能を左右する。

議システム“HITVISUAL1500”のブロック図を図7に示す。HITVISUAL1500は、1.5 Mビット/sの回線を使用したときに最も効率がよいように設計されているため、HV-100に比べて装置規模はやや大きくなっている。

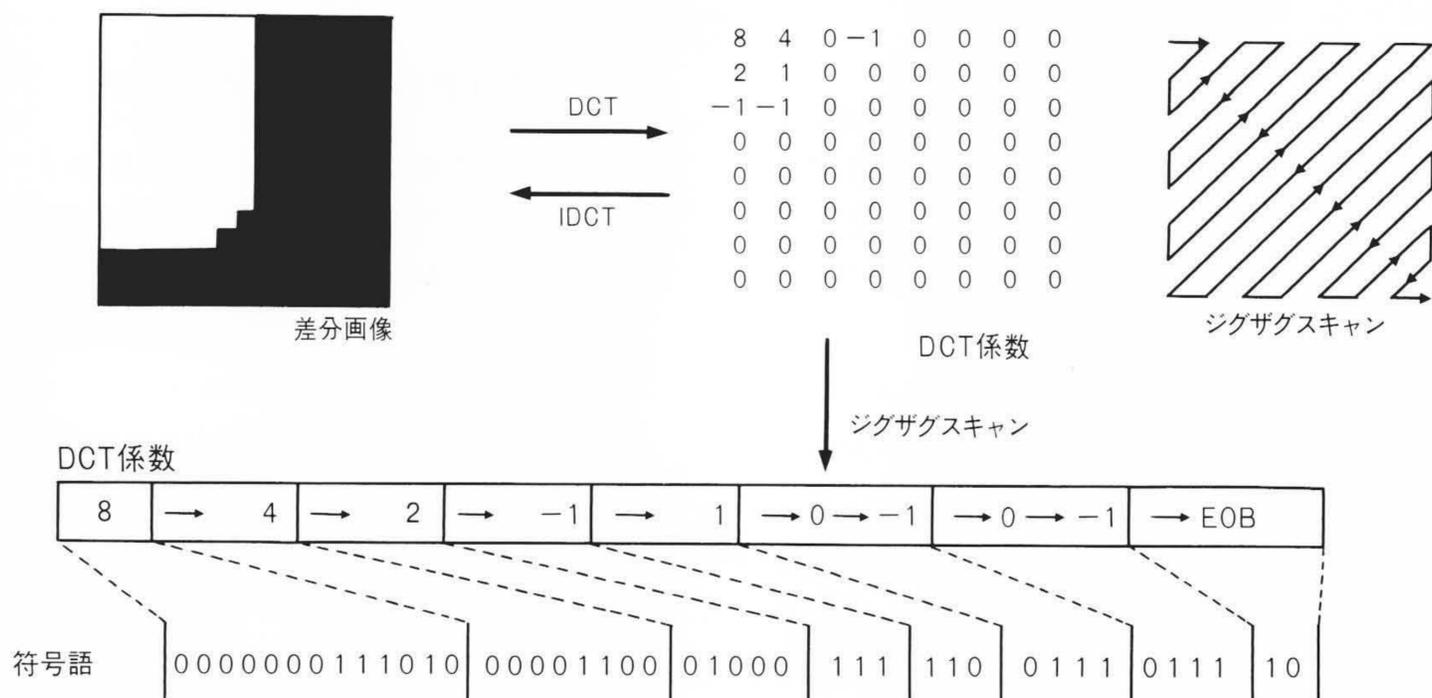
## 6 JPEGの符号化技術

JPEG(Joint Photographic Coding Experts Group)は静止画像の符号化方式を制定するためにISO(国際標準化機構)とCCITTの合同作業による委員会の名称であるが、一般には、この委員会の定めた静止画符号化の標準化を表す。

JPEGの基本アルゴリズムは先に述べたDCTと可変長符号化である。JPEGには以下の柔軟性を持つ特徴がある。

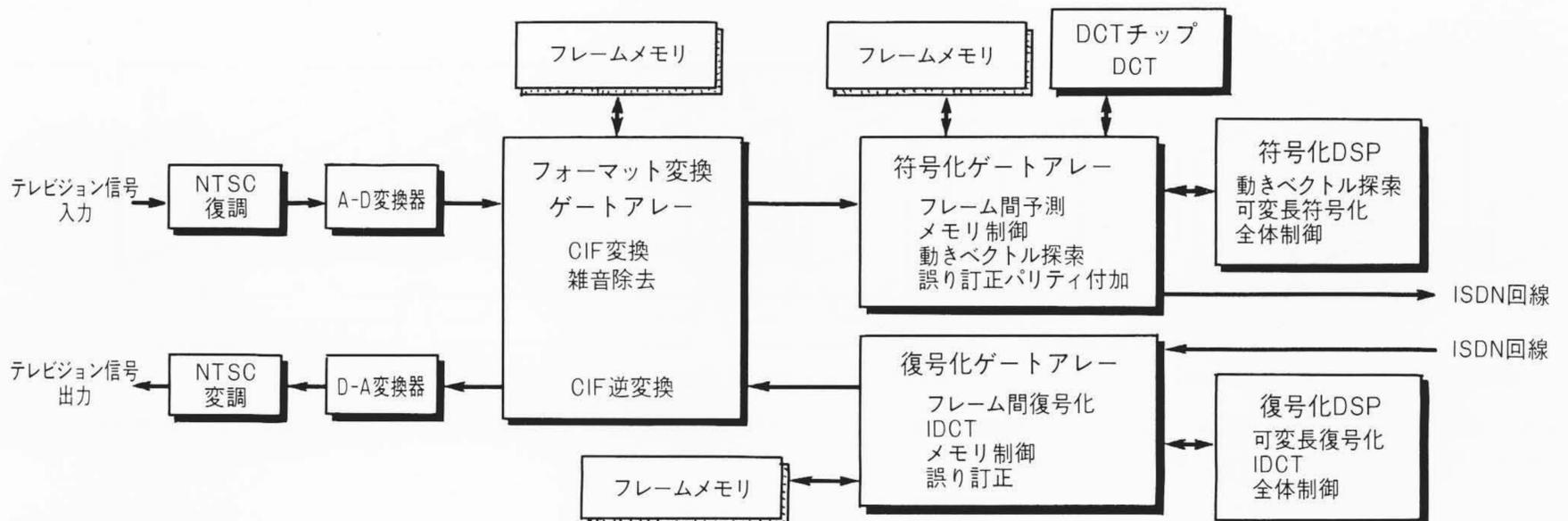
### (1) 解像度

テレビジョン信号程度の解像度(768×480程度)から医



注：略語説明 EOB (End of Block)

図4 DCTによる画像の符号化 低周波成分によってジグザグスキャンしたDCT係数に対し、連続するゼロの個数と非ゼロのDCT係数の値をセットにして可変長符号化する。



注：略語説明 CIF (Common Intermediate Format), DSP (Digital Signal Processor), NTSC (National Television System Committee), ISDN (Integrated Services Digital Network)

図6 VH-100のブロック図 小形化・経済化を図るためにH.261処理を徹底的に分析し、ハードウェアとソフトウェアの分担の最適化を行った。すべての機能がA4サイズ基板1枚に収められている。

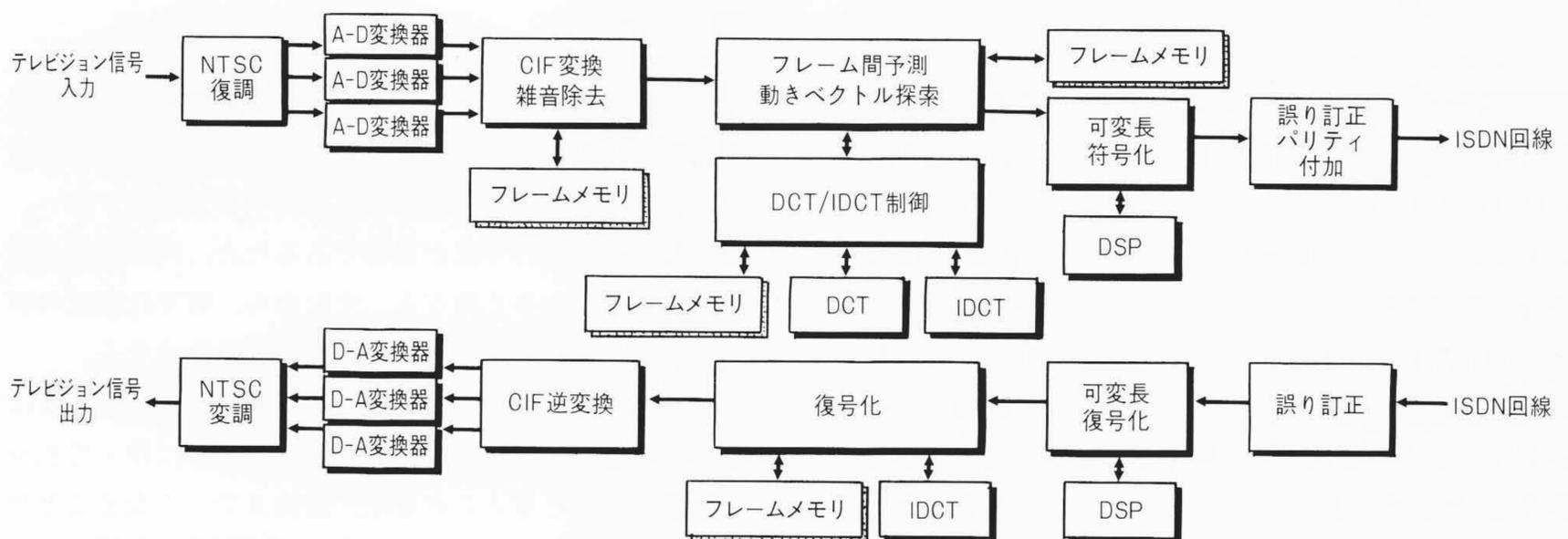


図7 HITVISUAL1500のブロック図 1.5 Mビット/sの伝送速度で最大限の性能を引き出すためのアーキテクチャを採用した。HV-100用のLSIを利用することにより、装置の小形化と経済化を同時に達成している。

療や印刷で用いられる超高解像度(10,000×10,000程度)までのさまざまな解像度に対応する。

(2) 可逆性

再生した画像と原画が同一である可逆符号化と、多少の劣化を許容して高圧縮する非可逆符号化がある。

(3) 階層性

画像を一時に伝送するシーケンシャル符号化と、低周波成分から順次伝送するプログレッシブ符号化がある。後者では粗い画像を先に表示し、順次細かい画像を表示できるため、受信者の心理的負担を軽減できる。

(4) DCT・DPCM

DCTの他にハード量の少ないDPCM(Differential Pulse Code Modulation)が使用できる。

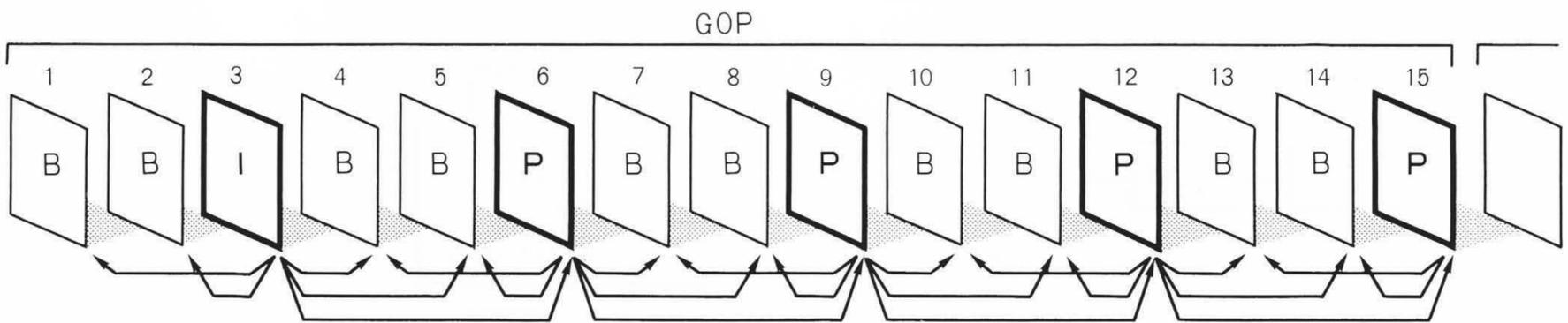
(5) 可変長符号化・算術符号化

可変長符号化(ハフマン符号化)と、圧縮性能の優れた算術符号化を選択できる。

通常JPEGというと、「DCT+ハフマン符号化」を示すことが多い。このほかの応用例としては、可逆符号化が必須(す)な場合はDPCM符号化、高圧縮率と高機能の両立を求める場合には「DCT+プログレッシブ符号化+算術符号化」などがある。

7 MPEG符号化方式

MPEG(Moving Picture Experts Group)は、CD-ROMなどの蓄積系メディアのための符号化方式である。蓄積系の符号化では、高画質、高速再生、ランダムアク



注：略語説明 GOP (Group of Picture), I (Intra picture), P (Predicted picture), B (Bidirectionally predicted picture)

図8 MPEGにおける予測構造 I, B, Pの3種類の画面から成る。Iピクチャは0.5秒に1回程度挿入され、特殊再生などに利用される。Iピクチャ、Pピクチャだけを見れば、H.261と同じ構造である。

セスなどが要求される。また、符号化装置は装置規模、価格などよりも符号化性能が要求され、逆に復号化装置は小形・経済性が要求される。MPEGではH.261に定期的フレーム内符号化、双方向予測、半画素単位の動き補償などを加えて上記の要求を満たしている。

MPEGでは、高速再生やランダムアクセスなどの特殊再生が可能のように、図8のようなGOP(Group of Picture)単位の構成をとっている。GOPはフレーム内符号化されるIピクチャ、H.261と同じ予測符号化されるPピクチャ、前後のIピクチャあるいはPピクチャから双方向予測符号化されるPピクチャから成る。双方向予測では、前のフレームか、後のフレームか、あるいは前後のフレームの平均か、いずれから予測するかを選択できる。特殊再生はIピクチャあるいはIピクチャとPピクチャを用いて行う。符号化順序は画像の入力順序とは異なり、Iピクチャを直前のBピクチャより先に符号化する。同図では「3→1→2→6→4→5…」の順に符号化される。

画面フォーマットはSIFと呼ばれ、NTSC圏では352×240ライン30フレーム/sである。

動きベクトルは0.5画素単位で指定ができる。整数の

ベクトルのときはH.261と同じ動作をするが、0.5画素のベクトルのときは、上下あるいは左右の画素から補間して新たな予測画素を作り出す。

## 8 おわりに

画像信号は多くの冗長度を含んでいるために、情報量の大幅な圧縮が可能である。また、符号化の方法も各種の標準化で決まり、今後急速に普及する可能性が高い。しかし、符号化の手法が複雑であるため、内部構成も設計者によって大きく異なる。すなわち、符号化装置のアーキテクチャが価格や装置規模を大きく左右する。

こういった、アーキテクチャが価値・装置規模に及ぼす影響の度合いは、今後のLSI技術の進歩に伴って徐々に小さくなると考えられるが、最後までなくなることはない。日立製作所では、このアーキテクチャの検討や符号化アルゴリズムの改良からLSI設計・ユーザーインタフェース設計まで、幅広い画像通信技術に対する検討を重ね、小形・低価格で、だれもが使いやすい画像通信機器を提供していく予定である。

## 参考文献

- 1) 木村, 外: ハイブリッドアーキテクチャを用いた小型画像CODEC, 電子情報通信学会研究会資料, IE91-58 (1991)
- 2) 安田: マルチメディア符号化の国際標準, 丸善(1991)